MICROCOPY RESOLUTION TEST CHART
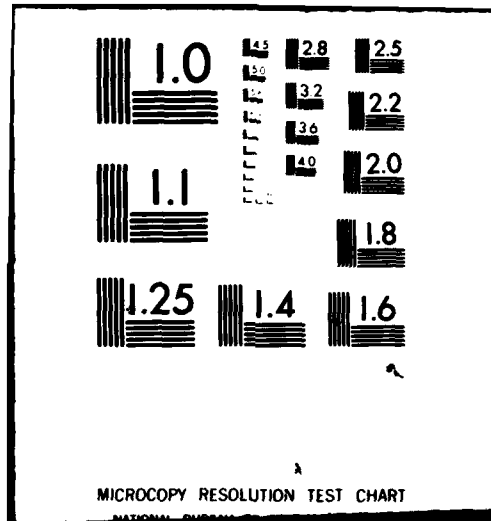
LEVEL

THE FISHER-BINGHAM DISTRIBUTION ON THE SPHERE

by

John T. Kent
Department of Statistics
University of Leeds
Leeds LS2 9JT U.K.

DTIC
SELEC
APR 1981

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

81 4   7 038

# The Fisher-Bingham distribution on the sphere

John T. Kent
Department of Statistics
University of Leeds
Leeds LS2 9JT U.K.

## Abstract

The Fisher distribution is the analogue on the sphere of the isotropic bivariate normal distribution in the plane. The purpose of this paper is to propose and analyze a spherical analogue of the general bivariate normal distribution. Estimation, hypothesis testing and confidence regions are also discussed.

## Introduction.

The _Fisher_ distribution on the sphere is the analogue of the _isotropic_ bivariate normal, that is the Fisher distribution has _circular contours_ of constant probability. However, in some problems it is desirable to have a more general distribution on the sphere with _oval contours_ in order to provide an analogue of the _general_ bivariate normal distribution. The purpose of this paper is to construct a suitable spherical analogue of the general bivariate normal distribution (denoted as the $FB_5$ distribution below).

After setting up our notation in Section 2, we define the 8-parameter Fisher-Bingham distribution $(FB_8)$ in Section 3. The $FB_5$ distribution will appear as a 5-parameter sub-family of $FB_8$. The limiting normal behaviour of $FB_5$ for large concentration and other motivating properties of $FB_5$ are discussed in Sections 4 and 6.

Sufficient statistics for the $FB_5$ distribution are described in Section 5, estimation of the parameters in Section 8, and a confidence region for the mean direction in Section 10. Several hypothesis tests of interest are discussed in Section 9.

An example to illustrate the use of the use of the $FB_5$ distribution is given in Section 11. Analogues of the Fisher-Bingham distribution in other dimensions are briefly mentioned in Section 12.

Although the primary emphasis in this paper is on the $FB_5$ distribution, properties which applying to other sub-families of $FB_8$ will also be mentioned where relevant.

## 2. Notation

Let $\Omega_3 = \{x \in R^3 : x'x = x_1^2 + x_2^2 + x_3^2 = 1\}$ denote the unit sphere in $R^3$. We can write $x$ in _polar coordinates_ $(\theta, \phi)$ defined by

$$x_1 = \cos\theta, \quad x_2 = \sin\theta\cos\phi, \quad x_3 = \sin\theta\sin\phi \tag{2.1}$$

where $0 \le \theta \le \pi$, $0 \le \phi < 2\pi$. If $dx$ denotes Lebesgue measure on $\Omega_3$, then in polar coordinates $dx = \sin\theta \, d\theta \, d\phi$. Throughout this paper we shall define distributions on $\Omega_3$ in terms of densities with respect to $dx$.

A useful way to plot spherical data is given by _Lambert's equal area projection_ (see e.g. Mardia, 1972, p.215) defined by

$$z_2 = \rho\cos\phi, \quad z_3 = \rho\sin\phi \tag{2.2}$$

where $\rho = 2\sin(\theta/2)$, $0 < \rho < 2$.

For any matrix $A(n \times p)$, let $A'$ denote the _transpose_ of $A$, $a_{(j)}(n \times 1)$ the $j$th _column_ of $A$, $j = 1, \ldots, p$, and $a_i(p \times 1)$ the $i$th _row_ of $A$ (written as a column vector), $i = 1, \ldots, n$.

An orthogonal matrix $\Gamma(3 \times 3)$ of positive determinant depends on 3 polar coordinates. Let us denote by $\Gamma = \Gamma(\psi, \eta, \xi)$ the matrix defined by

$$\Gamma = [g_{(1)}, \cos\xi \, g_{(2)} + \sin\xi \, g_{(3)}, -\sin\xi \, g_{(2)} + \cos\xi \, g_{(3)}] \tag{2.3}$$

where

$$G = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi\cos\eta & \cos\psi\cos\eta & -\sin\eta \\ \sin\psi\sin\eta & \cos\psi\sin\eta & \cos\eta \end{bmatrix} \tag{2.4}$$

and $0 \le \psi \le \pi$, $0 \le \eta, \xi < 2\pi$.

We next define a concept we shall need later in the paper. Given a non-zero vector $u(3 \times 1)$ and a symmetric matrix $A$, let

$E = E(u,A)$ be a (3×3) orthogonal matrix such that if $v = E'u$ and $B = E'AE$, then

$$v_2 = v_3 = 0, \quad b_{23} = 0, \tag{2.5}$$

$$v_1 > 0, \quad b_{22} \geq b_{33}. \tag{2.6}$$

Call the columns of $E$ the _constrained eigenvectors_ of $(u,A)$. Note that $e_{(1)}$ is proportional to $u$ whereas $e_{(2)}$ and $e_{(3)}$ diagonalize $A$ "as much as possible" subject to being constrained by $e_{(1)}$. Also note that $e_{(1)}$ defines a _vector_ (whose sign is determined by (2.6)), whereas $e_{(2)}$ and $e_{(3)}$ only define _axes_ (whose order is determined by (2.6)). The constrained eigenvectors can also be viewed as the eigenvectors after projecting $A$ onto the subspace orthogonal to $u$ (see Kato, 1966, pp 61-62).

It is also convenient to summarize $(u,A)$ in terms of the _size_ and _shape_ quantities,

$$r_1 = v_1 = \|u\| \quad \text{and} \quad r_2 = b_{22} - b_{33}, \tag{2.7}$$

respectively, which are invariant under orthogonal changes of the coordinate system.

For computational purposes, the matrix $E$ is most easily obtained by the following two-step procedure. First choose an orthogonal matrix $H$ to rotate $u$ to the north pole $(1,0,0)'$. (In the polar coordinates of (2.1) with $H$ for $\Gamma$, choose $\psi$ and $\eta$ so that $h_{(1)} = u$; here $\zeta$ is arbitrary, so for simplicity we can take $\zeta = 0$.) Then set $v = H'u$ and $C = H'AH$. Secondly, choose a rotation $K$ about the north pole to diagonalize $C_L$, where

$$C_L = \begin{bmatrix} c_{22} & c_{23} \\ c_{32} & c_{33} \end{bmatrix}$$

is the lower (2×2) submatrix of $C$. (In the polar coordinates of

(2.1) with $\underset{\sim}{K}$ for $\underset{\sim}{\Gamma}$, take $\psi = 0$, $\eta = 0$ and choose $\zeta$ to satisfy

$$\tan 2\zeta = 2c_{23}/(c_{22}-c_{33}),$$

ensuring that (2.6) also holds.) Then $\underset{\sim}{E} = \underset{\sim}{HK}$.

Note that even after the first stage the size and shape have simple interpretations in terms of $\underset{\sim}{w}$ and $\underset{\sim}{C}$. Letting $\ell_1$ and $\ell_2$ denote the eigenvalues of $\underset{\sim}{C}_L$ we have

$$r_1 = w_1, \quad r_2 = \ell_1 - \ell_2 . \tag{2.8}$$

## 3. The Fisher-Bingham distribution.

Define a distribution on $\underset{\sim}{\Omega}_3$ by the density

$$f(x) = \exp\{\kappa \underset{\sim}{v}'\underset{\sim}{x} + \sum_{j=2}^{3} \beta_j (\gamma'_{(j)}\underset{\sim}{x})^2\} , \quad \underset{\sim}{x}'\underset{\sim}{x} = 1. \tag{3.1}$$

We shall call (3.1) the Fisher-Bingham distribution since the first factor is proportional to a Fisher density and the second to a Bingham density. The 8 parameters of (3.1) are $\kappa \geq 0$, real-valued $\beta_2 \geq \beta_3$, a unit vector $\underset{\sim}{v}$, and an orthogonal matrix $\underset{\sim}{\Gamma} = [\underset{\sim}{\gamma}_{(1)}, \underset{\sim}{\gamma}_{(2)}, \underset{\sim}{\gamma}_{(3)}]$.

We shall also use the name $FB_8$ for the full family (3.1). Note that the constraint $\sum_1^3 x_i^2 = \sum_1^3 (\gamma'_{(i)}\underset{\sim}{x})^2 = 1$ implies that a term $\beta_1 (\gamma'_{(1)}\underset{\sim}{x})^2$ in the exponent of (3.1) would be redundant in the specification of the density.

The family of $FB_8$ distributions is closed under orthogonal transformations. If $\underset{\sim}{x}$ is a random vector from $FB_8(\kappa, \beta_2, \beta_3, \underset{\sim}{v}, \underset{\sim}{\Gamma})$ and $\underset{\sim}{H}$ is orthogonal, then $\underset{\sim}{H}'\underset{\sim}{x}$ comes from $FB_8(\kappa, \beta_2, \beta_3, \underset{\sim}{H}'\underset{\sim}{v}, \underset{\sim}{H}'\underset{\sim}{\Gamma})$. Note that the transformation $\underset{\sim}{x} \mapsto \underset{\sim}{H}'\underset{\sim}{x}$ can be thought of as changing the frame of reference, with the coordinate axes in the new frame given by the columns of $\underset{\sim}{H}$.

Several interesting distributions appear as special cases
of $FB_8$. Besides the uniform, Fisher and Bingham distributions
themselves, we also have the following families, all of which are
also closed under orthogonal transformations.

(a) $FB_6$ $(\kappa, \beta_2, \beta_3, \Gamma)$. Put $\underline{\nu} = \underline{\gamma}_{(1)}$ so that the Fisher axis
lines up with one of the Bingham axes. Then the number of
parameters is reduced by 2 to 6 parameters.

(b) $FB_5$ $(\kappa, \beta, \Gamma)$. Put $\underline{\nu} = \underline{\gamma}_{(1)}$ and set $\beta_2 = -\beta_3 = \beta$ say, with
$\beta \geq 0$. This distribution is a 5-parameter sub-family of $FB_6$ and
is proposed here as a spherical analogue of the bivariate normal
distribution. The justification for this proposal will be given
in Section 6.

(c) $FB_4$ $(\kappa, \beta, \underline{\gamma}_{(1)})$. Put $\underline{\nu} = \underline{\gamma}_{(1)}$ and set $\beta_2 = \beta_3 = \beta$, say.
Again, $FB_4$ is a sub-family of $FB_6$ but with very different behaviour
from $FB_5$. Note that since we cannot distinguish between $\underline{\gamma}_{(2)}$ and
$\underline{\gamma}_{(3)}$ here, $FB_4$ is only a 4-parameter family. If $\beta > \frac{1}{2}\kappa$, then
the mode of the $FB_4$ density is a small circle whose center lies
along the $\underline{\gamma}_{(1)}$ axis. This distribution was introduced and studied
by Bingham and Mardia (1978).

All of the above families of distributions are closed under
arbitrary rotations of the coordinate system. The inclusion
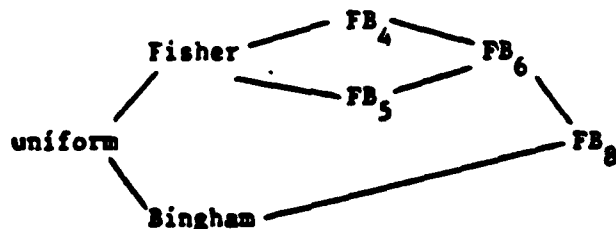relationships between them are summarized in Figure 1.



Figure 1.

The Fisher-Bingham distribution was first introduced in
Mardia (1975, p.352) on the sphere $\Omega_p, p \geq 2$. It also forms one
of a hierarchy of distributions considered in Beran (1979).
From Beran's point of view, the Fisher distribution contains an
arbitrary <u>linear</u> function of $\underset{\sim}{x}$ in the exponent of the density,
and $FB_8$ contains an arbitrary <u>linear</u> and <u>quadratic</u> function of
$\underset{\sim}{x}$. The other distributions in the hierarchy include higher order
polynomials in the exponent of the density. An extension of $FB_8$ to a
Stiefel manifold was proposed in Mardia and Khatri (1977).

As noted in Mardia and Khatri (1975), the $FB_8$ distribution
can be obtained by conditioning a trivariate normal distribution
with arbitrary mean vector and covariance matrix (a 9-parameter
family) to lie on the unit sphere.

Unfortunately, statistical work with the full $FB_8$ distribution
has been hampered by difficulties in estimating and interpreting
the parameters (but see de Waal, 1979). However, as we show in
this paper, these difficulties do not apply to the $FB_5$ distribution.
For $FB_5$ the parameters have important and natural interpretations,
and estimation is quite feasible.

4. <u>Limiting behaviour of $FB_6$ for large concentration.</u>

When the $FB_6$ distribution is highly concentrated about a point,
it is well-approximated by a bivariate normal distribution. This
result generalizes the well-known property of the Fisher distribution
(see e.g. Mardia, 1972, p.246), where an <u>isotropic</u> bivariate normal
appears. Details about the closeness of this approximation in the
Fisher case can be found in Kent (1978).

For convenience suppose that the orientation matrix $\underset{\sim}{\Gamma}$ equals

I, the identify matrix. Then in the polar coordinates (2.1), the $FB_6$ density in (3.1) takes the form

$$g(\theta,\phi) = \exp\{\kappa \cos \theta + \beta_2 \sin^2\theta \cos^2\phi + \beta_3 \sin^2\theta \sin^2\phi\} . \quad (4.1)$$

Theorem 4.1. Let $\underline{x} \sim FB_6 (\kappa, \beta_1, \beta_2, I)$ and let the parameters $\kappa, \beta_1, \beta_2$ vary in such a way that

$$\kappa \to \infty, \beta_2/\kappa \to d_2, \beta_3/\kappa \to d_3 \quad \text{with} \quad -\infty < d_3 \leq d_2 < \tfrac{1}{2} . \quad (4.2)$$

Then as $\kappa \to \infty$,

$$\kappa^{\frac{1}{2}} \begin{bmatrix} \theta \cos \phi \\ \theta \sin \phi \end{bmatrix} \doteq \kappa^{\frac{1}{2}} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} \xrightarrow{D} N_2 \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} , \begin{bmatrix} (1-2d_2)^{-1} & 0 \\ 0 & (1-2d_3)^{-1} \end{bmatrix} \right) . $$

$$(4.3)$$

where $\doteq$ denotes asymptotically equal in distribution and $\xrightarrow{D}$ denotes convergence in distribution.

Proof. Using the Taylor series expansions

$$\cos \theta = 1 - \theta^2/2 + \cdots , \quad \sin\theta = \theta + \cdots \quad (4.4)$$

for $\theta$ small and using (4.2) we see that (4.1) is approximately proportional to

$$\exp\{-\tfrac{1}{2}\kappa[\theta^2 - 2d_2\theta^2 \cos^2\phi - 2d_3\theta^2\sin^2\phi]\} , \quad (4.5)$$

which is the form of the limiting density in (4.3). To make this argument rigorous, it is merely necessary to show that

(a) the approximation (4.5) is adequate for $|\theta| < \kappa^{-\frac{1}{2}}$ and

(b) the probability mass associated with $|\theta| > \kappa^{-\frac{1}{2}}$ is negligible. The details are straightforward.

Similarly, it is straightforward to show that the two expressions on the left-hand side of (4.3) are asymptotically equal in distribution. □

<u>Corollary 4.1.</u>    Let   $\underset{\sim}{x} \sim FB_5(\kappa, \delta, I)$.   If   $\kappa$   and   $\beta$   vary in

such a way that

$$\kappa \to \infty, \qquad \delta/\kappa \to d \quad \text{with} \quad 0 \leq d < \tfrac{1}{2},$$

then

$$\kappa^{\frac{1}{2}} \begin{bmatrix} \theta \cos \bar{\phi} \\ \theta \sin \phi \end{bmatrix} \; \hat{=} \; \kappa^{\frac{1}{2}} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} \; \overset{D}{\to} \; N_2 \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} (1-2d)^{-1} & 0 \\ 0 & (1+2d)^{-1} \end{bmatrix} \right) .$$

## 5.  <u>Sufficient statistics</u>

The $FB_8$ distribution forms a canonical exponential family in

its 8 parameters.  If we write   $\underset{\sim}{\alpha} = \kappa \nu$   and   $\underset{\sim}{\Pi} = \Gamma \, \mathrm{diag}(0, \epsilon_2, \epsilon_3) \Gamma'$

then

$$f(x) \propto \exp\{\underset{\sim}{\alpha}'\underset{\sim}{x} + \underset{\sim}{x}'\underset{\sim}{\Pi}\underset{\sim}{x}\} \tag{5.1}$$

and a possible choice for the natural parameter vector is

$$(\alpha_1, \alpha_2, \alpha_3, \; \pi_{12}, \pi_{13}, \pi_{22}, \pi_{23}, \pi_{33})' \quad . \tag{5.2}$$

Given an   $(n \times 3)$   data matrix   $\underset{\sim}{X}$   from the $FB_8$ distribution,

the corresponding sufficient statistic is

$$\underset{\sim}{t} = \frac{1}{n} \sum_{i=1}^{n} (x_{i1}, x_{i2}, x_{i3}, \; 2x_{i1}x_{i2}, \; 2x_{i1}x_{i3}, \; x_{i2}^2, \; 2x_{i2}x_{i3}, \; x_{i3}^2)'$$

$$\tag{5.3}$$

Note that   $\underset{\sim}{t}$   holds the information contained in the <u>sample mean vector</u>

and the <u>sample dispersion matrix about</u>   0,

$$\underset{\sim}{\bar{x}} = \frac{1}{n} \sum \underset{\sim}{x}_i, \quad \underset{\sim}{S} = \frac{1}{n} \sum \underset{\sim}{x}_i \underset{\sim}{x}_i' \quad . \tag{5.4}$$

Unfortunately, the other families described above $(FB_4, FB_5$ and $FB_6)$

are <u>not</u> canonical exponential families, but instead from <u>curved</u>

exponential families.  In each case the parameter space has fewer

than 8 dimensions, but the minimal sufficient statistic is still

given by (5.3).

These remarks about $FB_4$, $FB_5$ and $FB_6$ apply only when all the parameters are unknown. A neater situation arises if we suppose that the mean direction $\underset{\sim}{\nu}$ (and also possibly the concentration $\kappa$) is known. With this knowledge $FB_4$, $FB_5$ and $FB_6$ (and also $FB_6$) now become canonical exponential families.

For definiteness we consider $FB_5$. After rotating the coordinate system so that $\underset{\sim}{\nu} = \underset{\sim}{\gamma}_{(1)}$ becomes equal to $(1,0,0)'$ the density is proportional to

$$f(\underset{\sim}{x}) = \exp\{\kappa x_1 + \delta_1 (x_2 - x_3)^2 + 2\ell_2 \, x_{23}^2\} , \qquad (5.5)$$

or in polar coordinates,

$$g(\theta, \phi) \propto \exp\{\kappa \cos\theta + \beta \sin^2\theta \, \cos 2(\phi - \lambda)\} , \qquad (5.6)$$

where

$$\delta_1 = \beta \cos 2\lambda , \qquad \delta_2 = \beta \sin 2\lambda . \qquad (5.7)$$

The parameter $\lambda \in [0, 2\pi)$ describes the direction of the major axis - see Section 6. Then the natural parameter and sufficient statistic are given by

$$(\kappa, \delta_1, \ell_2)' \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n}(x_{i1}, x_{i2}^2 - x_{i3}^2, 2x_{i2}x_{i3})' . \qquad (5.8)$$

If $\kappa$ is also known, the natural parameter and sufficient statistic become slightly simpler; namely

$$(\delta_1, \ell_2)' \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n}(x_{i2}^2 - x_{i3}^2, 2x_{i2}x_{i3})' . \qquad (5.9)$$

## 6. Properties of the $FB_5$ distribution.

The $FB_5$ density was defined in Section 3 by

$$f(\underset{\sim}{x}) = \exp\{\kappa \, \underset{\sim}{\gamma}_{(2)}'\underset{\sim}{x} + \beta[(\underset{\sim}{\gamma}_{(2)}'\underset{\sim}{x})^2 - (\underset{\sim}{\gamma}_{(3)}'\underset{\sim}{x})^2]\}.$$

As we shall see below the parameters can be described as follows: $\kappa \geq 0$ is the concentration, $\beta \geq 0$ describe the ovalness,

$Y_{(1)}$ is the mean direction or pole, $Y_{(2)}$ is the major axis, and $Y_{(3)}$ is the minor axis. Note that $Y_{(2)}$ and $Y_{(3)}$ are determined only up to sign, so that they do indeed define axes rather than directions.

If we rotate to the frame of reference defined by the columns of $\Gamma$, the density $f(x)$ takes a particularly simple form. For this reason we shall call this transformation, $x \mapsto x^* = \Gamma'x$, the transformation to the population standard frame of reference. The density for $x^*$ takes the form

$$f(x^*) = \exp\{\kappa x_1^* + \beta(x_2^{*2} - x_3^{*2})\}$$

or in polar coordinates

$$g(\theta,\phi) \propto \exp\{\kappa \cos\theta + \beta \sin^2\theta \cos 2\phi\} \ .$$

A sample analogue to the population standard frame will be defined in Section 8.

As stated in the introduction, the $FB_5$ distribution is proposed here as a spherical analogue of the bivariate normal distribution. The following properties show why this is a sensible proposal. We need to suppose here that $2\beta < \kappa$ to ensure the correct behaviour.

(a)   $FB_5$ and the bivariate normal are both 5-parameter families.

(b)   The contours of constant probability near the pole $Y_{(1)}$ are approximately ellipses with major and minor axes $Y_{(2)}$ and $Y_{(3)}$, respectively. (This property follows easily from (4.4).)

(c)   The geometric average of $g(\theta,\phi)$ over circles of constant latitude is proportional to a Fisher density, that is,

$$\int_0^{2\pi} \log g(\theta,\phi)d\phi = \kappa\cos\theta + \text{constant}.$$

Thus, in this sense $FB_5$ is a natural extension of the Fisher distribution.

(d)    As $\theta$ goes from $0$ to $\pi$ for fixed $\phi$, $g(\theta,\phi)$ decreases monotonically. Thus $g(\theta,\phi)$ is unimodal on all great circles through the pole.

(e)    For large values of the concentration parameter $\kappa$, $FB_5$ is approximately the same as a bivariate normal distribution with mean $\underline{\gamma}_{(1)}$ and major and minor axes $\underline{\gamma}_{(2)}$ and $\underline{\gamma}_{(3)}$ respectively. (See Corollary 4.1.)

Note that the larger $FB_6$ family is not a suitable spherical analogue of the general bivariate normal distribution because it has one too many parameters. In Theorem 4.1 we saw that this "extra" parameter is "asymptotically unidentifiable" for large concentration.

Hence we have introduced a further constraint $(\beta_2 = -\beta_3)$ to define the $FB_5$ family. To some extent this constraint is arbitrary (in fact a different constraint was proposed in Kent, 1980). However, the constraint used here does have some theoretically attractive properties ((c) and (d) above). Moreover, as we shall see in the next section, with this definition of $FB_5$ the normalization constant takes a reasonably tractable form.

## 7.  Moments of $FB_5$.

So far we have not dealt with the normalization constant of the $FB_5$ distribution,

$$c(\kappa,\beta) = \int_0^\pi \int_0^{2\pi} \exp\{\kappa\cos\theta + \beta\sin^2\theta\,\cos2\phi\}\sin\theta\,d\phi\,\,d\theta. \qquad (7.1)$$

Using the results

$$\int_0^{\pi/2} (\sin\theta)^a (\cos\theta)^b \, d\theta = \tfrac{1}{2} B(\tfrac{a+1}{2}, \tfrac{b+1}{2}), \qquad (7.2)$$

(Abramowitz and Stegun, 1972, (6.2.1), p.258) where $B(\cdot,\cdot)$ is the beta function, and

$$\int_0^\pi e^{\kappa\cos\theta} \sin^{2\nu}\theta \, d\theta = \pi^{\tfrac{1}{2}} \Gamma(\nu+\tfrac{1}{2})(\tfrac{1}{2}\kappa)^{-\nu} I_\nu(\kappa) \qquad (7.3)$$

(Abramowitz and Stegun, 1972, (9.6.18), p.376) where $I_\nu(\kappa)$ is the modified Bessel function, we can expand $c(\kappa,\beta)$ in a series

$$c(\kappa,\beta) = 2\pi \sum_{k=0}^\infty \frac{\Gamma(k+\tfrac{1}{2})}{\Gamma(k+1)} \beta^{2k} (\tfrac{1}{2}\kappa)^{-2k-\tfrac{1}{2}} I_{2k+\tfrac{1}{2}}(\kappa) . \qquad (7.4)$$

Since sequences of Bessel functions can be quickly and easily computed by the method of Amos (1974), formula (7.4) provides a quick and simple method of calculating $c(\kappa,\beta)$. Note that $c(o,o) = 4\pi$, the surface area of the sphere, and

$$c(\kappa,o) = 4\pi\kappa^{-1} \sinh\kappa ,$$

the normalizing constant for the Fisher distribution.

For large $\kappa$ (with $2\beta/\kappa<1$ fixed) we have from Corollary 4.1 the asymptotic formula

$$c(\kappa,\beta) \stackrel{.}{=} 2\pi e^\kappa [(\kappa-2\beta)(\kappa+2\beta)]^{-1/2} . \qquad (7.5)$$

Consider a random vector $\underline{x} \sim FB_5(\kappa,\beta,I)$. Differentiating (7.1) under the integral sign and writing $c = c(\kappa,\beta)$, $c_\kappa = \partial c(\kappa,\beta)/\partial\kappa$, etc., we find

$$E x_1 = c_\kappa/c \ , \quad E x_1^2 = c_{\kappa\kappa}/c$$
$$E(x_2^2 - x_3^2) = c_\beta/c \ , \qquad\qquad (7.6)$$

and hence since $x_1^2 + x_2^2 + x_3^2 = 1$,

$$E x_2^2 = (c - c_{\kappa\kappa} + c_\beta)/2c$$

$$E x_3^2 = (c - c_{\kappa\kappa} - c_\beta)/2c. \tag{7.7}$$

For later use write

$$\mu = E x_1 \quad \text{and} \quad \sigma_j^2 = E x_j^2, \quad j = 1,2,3. \tag{7.8}$$

By symmetry, most of the other moments of interest equal 0,

$$E(x_2) = E(x_3) = 0, \tag{7.9}$$

$$E(x_2 x_3) = 0, \tag{7.10}$$

$$E(x_1 x_2) = E(x_1 x_3) = 0. \tag{7.11}$$

## 8. Moment Estimation for $FB_5$.

Let $\underline{x}_1, \ldots, \underline{x}_n$ be a sample from $FB_5(\kappa, \beta, \Gamma)$. The standard way to estimate the parameters of $FB_5$ is to use maximum likelihood estimates $\hat\kappa, \hat\beta, \hat\Gamma$. However, it does not seem possible to obtain explicit expressions for the m.l.e.s, so iterative methods must be used to find them.

In this section we propose simpler estimates which we call the moment estimates $\bar\kappa, \bar\beta, \bar\Gamma$ for the parameters of $FB_5$. They have the following properties.

(a)  The moment estimates are consistent estimates of the true parameters and hence provide suitable starting values for maximum likelihood iteration.

(b)  The orientation matrix $\bar\Gamma$ can be calculated explicitly.

(c)  If either the eccentricity $2\beta/\kappa$ is small (the usual case in practice) or if $\kappa$ is large, then the moment estimates are close to the m.l.e.s.

(d)  If the data is highly concentrated, the concentration parameters $\bar\kappa, \bar\beta$ can also be calculated explicitly.

More specifically the moment estimates are defined as follows. Let $\bar{x}$ and $S$ be the sample mean vector and the sample dispersion matrix about $0$ as in (5.4). Then, in the terminology of Section 2, $\hat{\Gamma}$ is defined to be the matrix of constrained eigenvectors for $(\bar{x}, S)$. Further, letting $r_1 = \|\bar{x}\|$ and $r_2$ denote the size and shape quantities for $(\bar{x}, S)$, the estimates $\hat{\kappa}, \hat{\beta}$ of concentration parameters are determined implicitly by the equations

$$r_1 - c_\kappa/c = 0, \quad r_2 - c_\beta/c = 0. \tag{6.1}$$

For large $\kappa$, the use of (7.5) leads to the explicit solution

$$\hat{\kappa} = (2 - 2r_1 - r_2)^{-1} + (2 - 2r_1 + r_2)^{-1}$$

$$\hat{\beta} = \tfrac{1}{4}\{(2 - 2r_1 - r_2)^{-1} - (2 - 2r_1 + r_2)^{-1}\}. \tag{8.2}$$

The orientation matrix $\hat{\Gamma}$ has been chosen so that for $x^*$ the sample analogues of (7.9)-(7.10) (but not (7.11)) will hold. For this reason we shall term the transformation $x \mapsto x^* = \hat{\Gamma}'x$ the transformation to the sample standard frame of reference.

The rationale behind the moment estimate $\hat{\Gamma}$ is as follows. The first column $\hat{\gamma}_{(1)}$ is the mean direction of the sample, which is also the m.l.e. of the mean direction under a Fisher distribution. If the eccentricity $2\beta/\kappa$ is not too large (which is the most important case in practice), then $\hat{\gamma}_{(1)}$ will also be close to the m.l.e. of the mean direction for $FB_5$. Further, if the true mean direction $\hat{\gamma}_{(1)} = \gamma_{(1)}$ were known, then (7.3)-(7.4) would ensure that $\hat{\gamma}_{(2)}$ and $\hat{\gamma}_{(3)}$ would be the m.l.e.s of the major and minor axes, respectively. Hence if $2\beta/\kappa$ is not too large, the moment estimates should be nearly as efficient as the maximum likelihood estimates.

A similar situation arises for large $\kappa$ when $FB_5$ is close to a bivariate normal. Then the moment estimates and m.l.e.s for $FB_5$ will both be close to the corresponding m.l.e.s for the bivariate normal.

## 9. Some hypothesis tests.

In this section we describe several large-sample hypothesis tests of interest. All of these tests can be carried out using the following general result.

Theorem 9.1. Consider $n$ independent identically distributed observations from a model $H_1$ with parameters $(\tau, \lambda)$ of dimensions $p$ and $q$ respectively, and consider a null hypothesis $H_0 : \lambda = 0$. Suppose that under $H_0$, the model forms a canonical exponential family for $\tau$ with minimal sufficient statistic $u$, and that under $H_1$ (with $\tau$ known) the model forms a canonical exponential family for $\lambda$ with minimal sufficient statistic $v$. Define Rao's score statistic by

$$W_u = [v - E(v|u)]' \; Var(v|u)^{-1} \; [v - E(v|u)] , \qquad (9.1)$$

where all moments are calculated under $H_0$ with $\tau = \hat{\tau}$ ($\hat{\tau}$ being the m.l.e. of $\tau$ under $H_0$).

Then asymptotically as the sample size $n \to \infty$

$$W_u \sim \chi_q^2 , \qquad (9.2)$$

and further $W_u$ is asymptotically equivalent to the likelihood ratio statistic $-2 \log L$ for $H_0$ vs $H_1$. We reject $H_0$ if $W_u$ is too large.

Proof. This result is a special case of a general result in Cox and Hinkley (1974), p.324, equation after (5.6). See also Rao (1973), p.418.

$\square$

Note that the score statistic $W_u$ is usually simpler to calculate than the likelihood ratio statistic because only a parametric model under $H_0$ need be fitted. Some hypothesis tests which fit into this framework will now be described.

(a) $H_0$ : Fisher vs $H_1$ : $FB_5$

A Fisher distribution with concentration parameter $\kappa$ and mean direction $\underline{v}$ forms a canonical exponential family with natural parameters $(\kappa v_1, \kappa v_2, \kappa v_3)$ and sufficient statistic $\underline{u} = n^{-1} \sum \underline{x}_i$. The m.l.e.s satisfy

$$\hat{\underline{v}} = n^{-1} \sum \underline{x}_i / r_1 , \quad I_{3/2}(\hat{\kappa})/I_{1/2}(\hat{\kappa}) = r_1 \qquad (9.3)$$

where $r_1 = \|n^{-1} \sum \underline{x}_i\|$ is the resultant length, also used in (8.1). Now let $\underline{H}$ be an orthogonal matrix which depends on the data only through $\hat{\underline{v}}$, and whose first column is given by $\underline{h}_{(1)} = \hat{\underline{v}}$. Let $\underline{y}_i = \underline{H}'\underline{x}_i$, $i = 1, \ldots, n$. Then from (5.9), with $\kappa = \hat{\kappa}$ and $\underline{\gamma}_{(1)} = \hat{\underline{v}}$ assumed known, the model under $H_1$ forms a canonical exponential family with sufficient statistics

$$v_1 = \frac{1}{n} \sum (y_{i2}^2 - y_{i3}^2) , \quad v_2 = \frac{2}{n} \sum y_{i2} y_{i3} . \qquad (9.4)$$

Now the assumption that the $\underline{x}_i$ come from a Fisher distribution $F(\hat{\kappa}, \hat{\underline{v}})$ is equivalent to the assumption that the $\underline{y}_i$ come from $F(\hat{\kappa}, (1,0,0)')$. Further, for a random vector $\underline{y} \sim F(\hat{\kappa}, (1,0,0)')$, we have by symmetry

$$E(y_2^2 - y_3^2) = 0 , \quad E(y_2 y_3) = 0$$

$$E(y_2^3 y_3) = E(y_2 y_3^3) = 0$$

$$E(y_j y_2^2 - y_j y_3^2) = 0 , \quad E(y_j y_2 y_3) = 0 , \quad j = 1, 2, 3,$$

and we have by (7.2) and (7.3),

$$E(y_2^2 - y_3^2) = E(4 y_2^2 y_3^2) = (\hat{\kappa}/2)^{-2} \quad I_{5/2}(\hat{\kappa})/I_{1/2}(\hat{\kappa}) .$$

Hence the test statistic (9.1) takes the form

$$n \frac{(\hat{\kappa}/2)^2 I_{1/2}(\hat{\kappa})}{I_{5/2}(\hat{\kappa})} (v_1^2 + v_2^2) \sim \chi_2^2 . \qquad (9.5)$$

Using the notation of Section 2, it is easy to check that $v_1^2 + v_2^2$ equals the squared shape quantity $r_2^2$ for $(\bar{x}, S)$. In particular it is clear that the statistic $W_u$ does not depend on the arbitrariness in the choice of the second and third columns of $H$.

In practice this test might be used in the following situation. Given a set of spherical data, an experimenter might first look for directionality by testing $H_0$: uniform vs $H_1$: Fisher (the Rayleigh test). If this null hypothesis is rejected he might assess the circular symmetry of the data about the pole by using the test described here.

For large concentration, this test reduces to a test of sphericity for the bivariate normal distribution; see for example Mardia, Kent, and Bibby (1979) p. 134.

(b)     $H_0$: Fisher vs $H_1$: $FB_8$.

As in (a), this is a goodness-of-fit test for the Fisher distribution, but here with a more general alternative in mind. The calculations are similar to those in (a) but somewhat more involved. Details are given in Mardia and Holmes (1980).

The analogous test on the circle was given by Cox (1975); see also Cox and Barndorff-Nielsen (1979), p.291.


(c)   $H_0$: Bingham vs $H_1$: $FB_8$.

This test is included here because it fits the assumptions of Theorem 9.1, but the application of this test is somewhat different from (a) and (b). Suppose an experimenter has data to which he would like to fit a Bingham distribution. The data are suspected to be but not known to be antipodally symmetric. There are two ways to proceed in this situation.

(i)   First use a Rayleigh test ($3n\ \bar{x}'\bar{x} - \chi_3^2$) to test uniform vs

Fisher. If uniformity is accepted, then antipodal symmetry can be presumed and a Bingham distribution can be fitted.

(ii) First fit a Bingham distribution, and then assess the goodness-of-fit by using the test of this section. This latter approach seems more suitable when the data is known at the outset not to be uniformly distributed.

Under $H_o$, $S$, the sample dispersion matrix about $0$, is sufficient for the parameters. If $x$ comes from a Bingham distribution, we have by symmetry

$$E(x_i) = 0 , \quad E(x_i x_j x_k) = 0, \quad i,j,k = 1,2,3$$

so that

$$E(\bar{x}) = 0 , \quad E(\bar{x}|S) = 0 .$$

Further, since the m.l.e.s of the Bingham parameters are chosen so that $E(xx') = S$, we have

$$Var(\bar{x}|S) = Var(\bar{x}) = \frac{1}{n} S.$$ Hence the score statistic takes the form

$$n \bar{x}'S^{-1}\bar{x} - \chi_3^2 . \tag{9.6}$$

Note that when the data is uniformly distributed, $S = \frac{1}{3}I$, so that (9.6) reduces to the Rayleigh test statistic.

## 10. A confidence interval for the mean direction.

Let $x_1,\ldots,x_n$ be a sample from $FB_5(\kappa,\beta,\Gamma)$. First, transform to the population standard frame (see Section 6) $y_i = \Gamma'x_i, i = 1,\ldots,n$, so that the $y_i$ form a sample from $FB_5(\kappa,\beta,I)$. Then by the central limit theorem and (7.6)-(7.11) we see that the sample mean $\bar{y} = n^{-1}\sum y_i$ is asymptotically normally distributed with mean vector $(\mu,0,0)'$ and covariance matrix $n^{-1} diag(\sigma_1^2,\sigma_2^2,\sigma_3^2)$.

The sample mean direction for the $y_i$ (which is also the moment estimate of the true mean direction of the $y_i$) is defined by $\bar{y}_{(1)} = \bar{y}/\|\bar{y}\|$. Consider the two coordinates $(\gamma_{21}, \gamma_{31})'$ of $\bar{y}_{(1)}$. Then by a general result on transformations (see, e.g. Rao, 1973, p. 387), $(\bar{\gamma}_{21}, \bar{\gamma}_{31})$ is also asymptotically normally distributed,

$$\begin{bmatrix} \bar{\gamma}_{21} \\ \bar{\gamma}_{31} \end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} , \frac{1}{na^2} \begin{bmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_3^2 \end{bmatrix} \right) .$$

Hence, an asymptotic $100(1-\alpha)\%$ probability region about the mean direction of $FB_5(\kappa,\beta,\Gamma)$ is given in population standard coordinates by the ellipse-like region

$$\{\bar{\gamma}_{(1)} \in \Omega_3 : \bar{\gamma}_{11} > 0, \quad n \mu^2(\bar{\gamma}_{21}^2/\sigma_2^2 + \bar{\gamma}_{31}^2/\sigma_3^2) \leq \chi_{2;\alpha}^2 \} , \qquad (10.1)$$

where $\chi_{2;\alpha}^2$ denotes the upper $\alpha$ critical value of the $\chi_2^2$ distribution.

Inverting (10.1) provides a confidence region for the true mean direction about the sample mean direction. More specifically, suppose $x_1, \ldots, x_n$ come from $FB_5(\kappa,\beta,\Gamma)$ with true mean direction $\nu = \gamma_{(1)}$. Let $\Gamma$ be the moment estimate of $\Gamma$ (see Section 7) and transform to the <u>sample standard frame</u>, $x^*_i = \bar{\Gamma}' x_i$, $i = 1, \ldots, n$, and let $\nu^* = \bar{\Gamma}'\nu$. Then an asymptotic confidence region for $\nu^*$ is defined by the ellipse-like region on the sphere

$$\{\nu^* \in \Omega_3 : \nu_1^* > 0, \quad n \mu^2(\nu_2^{*2}/\sigma_2^2 + \nu_3^{*2}/\sigma_3^2) < \chi_{2;\alpha}^2 \} \qquad (10.2)$$

Of course, in practice $\mu, \sigma_2^2$ and $\sigma_3^2$ will not be known, but must be replaced by any consistent estimates. One possibility is to estimate $\kappa$ and $\beta$ and then use (7.6)-(7.7). However, a simpler technique is to just use the sample moments

$$\tilde{\mu} = n^{-1} \sum x_{i1}^* \ , \quad \tilde{c}_2^2 = n^{-1} \sum x_{i2}^{*2} \quad \text{and} \quad \tilde{\sigma}_3^2 = n^{-1} \sum x_{i3}^{*2} \ . \qquad (10.3)$$

respectively.

Note that in moving from (10.1) to (10.2) we have switched from a frame of reference about the _true_ mean direction to a frame about the _sample_ mean direction. However, since these two points lie within $O(n^{-1})$ of one another in probability, the complications arising from this switch are negligible and (10.2) remains asymptotically valid.

When using the equal-area projection (2.2) (in the sample standard frame) and when using the estimates (10.3), a confidence region asymptotically equivalent to (10.2) is given by

$$\{(z_2, z_3) : n \, \tilde{\mu}^2 (z_2^2/\tilde{\sigma}_2^2 + z_3^2/\tilde{c}_3^2) < \chi_{2;\alpha}^2 \}. \qquad (10.4)$$

## 11. Example.

Creer, Irving and Nairn (1959) measured directions of magnetism at $n = 34$ sites in the Great Whin Sill. Their data is summarized in Table 1, column (b) of that paper, pp.311-312 (excluding sites 32 and 34). Their use of declination (D) and inclination (I) is related to our use of polar coordinates in (2.1) by $\theta = 90° + I$, $\phi = 360° - D$. [Change program accordinly).

The summary statistics are given by

$$\bar{x} = \begin{bmatrix} 0.083 \\ -0.959 \\ -0.131 \end{bmatrix} , \quad S = \begin{bmatrix} 0.045 & -0.075 & 0.014 \\ -0.075 & 0.921 & -0.122 \\ 0.014 & -0.122 & 0.034 \end{bmatrix} ,$$

from which we find $r_1 = 0.971$, $r_2 = 0.0229$.

The moment and maximum likelihood estimates of location are both given to three decimal places by

$$\hat{\underline{\Gamma}} \approx \tilde{\underline{\Gamma}} = \begin{bmatrix} 0.085 & -0.979 & 0.185 \\ -0.987 & -0.108 & -0.117 \\ 0.134 & -0.172 & -0.976 \end{bmatrix} \quad ,$$

and the estimates of concentration by

$$\tilde{\kappa} = 42.16 \qquad \tilde{\delta} = 9.27 \qquad \text{(exact moment)}$$

$$\bar{\kappa} = 41.76 \qquad \bar{\beta} = 8.37 \qquad \text{(asymptotic moment)}$$

$$\bar{\kappa} = 42.16 \qquad \bar{\beta} = 9.28 \qquad \text{(maximum likelihood)}.$$

The data and a 95% confidence region for the mean direction based on (10.2) and (10.3) are given in Figure 2, plotted using the equal-area projection of (2.2) in the sample standard frame.

The hypothesis test of Section 9(b) and the corresponding likelihood ratio test yield the values

$$W_u = 5.96 \quad \text{and} \quad -2 \log L = 6.55.$$

Since the upper 5% critical value of $\chi_2^2$ is 5.99, both statistics show moderate evidence of a departure from a Fisher distribution.

In this example an alternative approach was used by Creer et.al. They managed to transform the data to approximate circular symmetry and then to use statistical techniques applicable to the Fisher distribution.

## 12. Analogues in other dimensions.

Much of the theory in this paper extends, at least in principle, to other dimensions. The analogue of the full Fisher Bingham family on the unit sphere in $R^p$, $p>2$, can be written in the general form (2.1) with the summation from 2 to 3 replaced by a summation from 2 to p (as in Beran, 1979).

An analogue of $FB_5$ can be obtained by introducing the constraint

$$\sum_{j=2}^{p} \beta_j = 0.$$

On the circle, $p = 2$, this analogue of $FB_5$ is no more general than the von Mises distribution itself. For general $p \geq 3$, provided $|\beta_j| < \kappa/2$, $j = 2, \ldots, p$, properties analogous to those in Sections 4-6 are valid. When $p > 3$, moment estimation still can be carried out, although the normalization constant seems to become more complicated as the dimension increases.
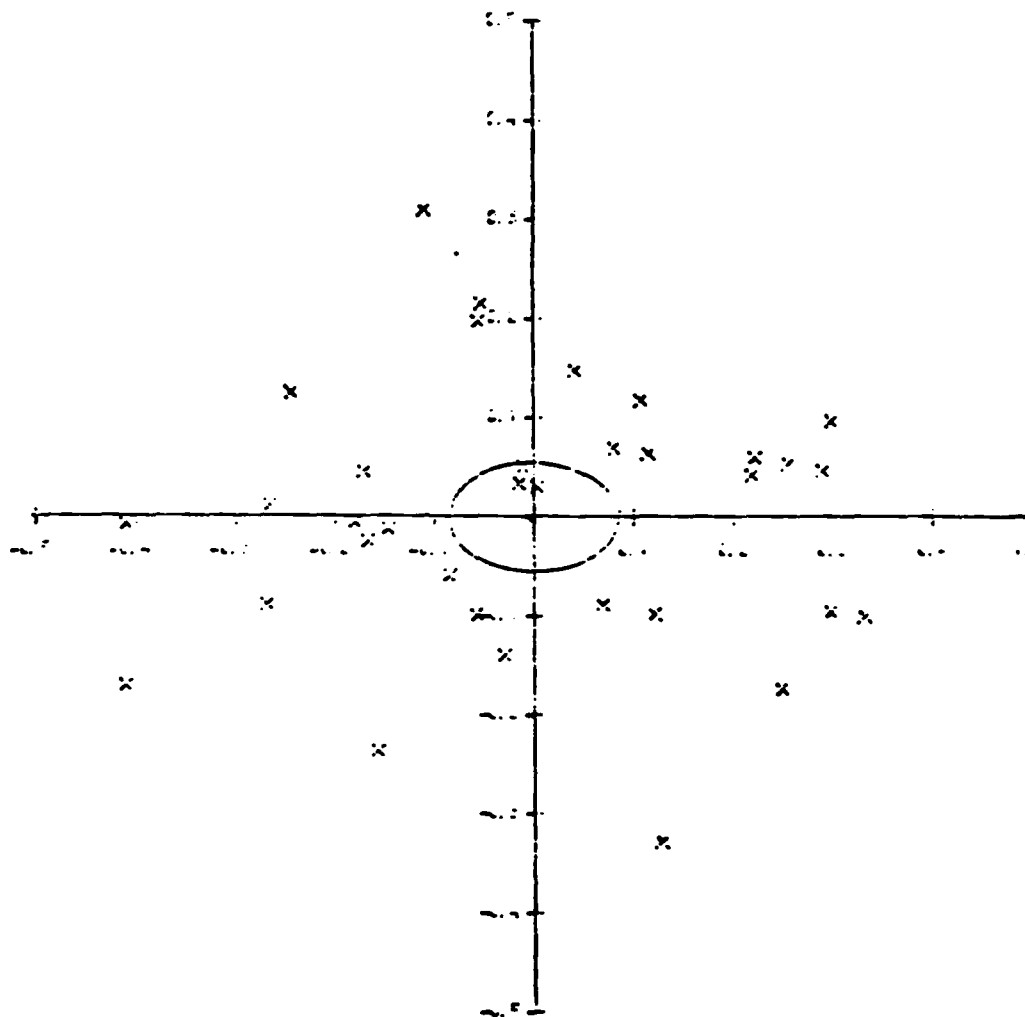
Figure 2:  Plot of the Great Whin Sill data using an equal-area projection in the sample standard frame, and a 95% confidence ellipse for the mean direction.

REFERENCES

1. Abramowitz, M. and Stegun, I. (1972) Handbook of Mathematical
   Functions. Dover, New York.

2. Amos, D.E. (1974) Computation of the modified Bessel functions
   and their ratios, Math. of Computation 28, 239-51.

3. Barndorff-Nielsen, O. and Cox, D.R. (1979) Edgeworth and saddle-point
   approximations with statistical applications (with Discussion),
   J.Royal Statist.Soc. B 41, 279-312.

4. Beran, R. (1979) Exponential models for directional data, Ann.Statist.
   7, 1162-78.

5. Bingham, C. and Mardia, K.V. (1978) A small circle distribution
   on the sphere, Biometrika 65, 379-89.

6. Cox, D.R. (1975) Discussion following Mardia, K.V. J.Royal Statist.Soc.
   B 37, 380-1.

7. Creer, K.M., Irving, E. and Nairn, A.E.M. (1959) Paleomagnetism of
   the Great Whin Sill, Geophysical J. of the Royal Astron.Soc.
   2, 306-323.

8. de Waal, D.J. (1979) On the normalizing constant for the Bingham-
   von Mises-Fisher matrix distribution. 13, 103-112.

9. Kato, T. (1966) Perturbation Theory for Linear Operators. Springer-
   Verlag, Berlin.

10. Kent, J.T. (1978) Limiting behaviour of the von Mises-Fisher
    distribution, Math.Proc.Camb.Phil.Soc.84, 531-536.

11. Kent, J.T. (1979) Discussion following Barndorff-Nielsen, O. and
    Cox, D.R. J.Royal Statist.Soc. B 41, 305-306.

12. Khatri, C.G. and Mardia, K.V. (1975) The von Mises-Fisher matrix
    distribution, Research Report No.1, Directional Data Analysis
    Project, Department of Statistics, University of Leeds.

13. Khatri, C.G. and Mardia, K.V. (1977)  The von Mises-Fisher matrix distribution in orientation statistics, <u>J.Royal Statist.Soc.</u> B <u>39</u>, 95-106.

14. Mardia, K.V. (1972)  <u>Statistics of Directional Data</u>. Academic Press, London.

15. Mardia, K.V. (1975)  Statistics of directional data (with discussion), <u>J. Royal Statist.Soc.</u> B <u>37</u>, 349-393.

16. Mardia, K.V. and Holmes, D. (1980)  A test of Fisherness (unpublished manuscript).

17. Mardia, K.V., Kent, J.T., and Bibby, J.M. (1979)  <u>Multivariate Analysis</u>. Academic Press, London.

18. Rao, C.R. (1973)  <u>Linear Statistical Inference and its Applications</u>, <u>second edition</u>. John Wiley and Sons, New York.

TR-181-SER-2

| REPORT DOCUMENTATION PAGE | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|

| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
|---|---|---|
| T.R.#181, Series 2 | AD A094475 | |

| 4. TITLE (and Subtitle) | 5. TYPE OF REPORT & PERIOD COVERED |
|---|---|
| THE FISHER-BINGHAM DISTRIBUTION ON THE SPHERE. | Technical report. |
| | 6. PERFORMING ORG. REPORT NUMBER |

| 7. AUTHOR(s) | 8. CONTRACT OR GRANT NUMBER(s) |
|---|---|
| John T. Kent | N00014-79-C-0322 |

| 9. PERFORMING ORGANIZATION NAME AND ADDRESS | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
|---|---|
| Department of Statistics Princeton University Princeton, N.J. 08540 | Dec 80 |

| 11. CONTROLLING OFFICE NAME AND ADDRESS | 12. REPORT DATE |
|---|---|
| Office of Naval Research (Code 436) Arlington, Virginia 22217 | |
| | 13. NUMBER OF PAGES |
| | 25 |

| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | 15. SECURITY CLASS. (of this report) |
|---|---|
| | UNCLASSIFIED |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

directional data, Fisher distribution, constrained eigenvectors

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Fisher distribution is the analogue on the sphere of the isotropic bivariate normal distribution in the plane. The purpose of this paper is to propose and analyze a spherical analogue of the general bivariate normal distribution. Estimation, hypothesis testing and confidence regions are also discussed.

406873